IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Appln No.:	09/692,846
Applicants:	Konopka, Courtney C.
Filed:	October 19, 2000
For:	NATURAL LANGUAGE INTERFACE CONTROL SYSTEM
TC/A.U.:	2654
Examiner:	Lamont Spooner

This document was filed electronically using the USPTO's EFS-WEB.

AMENDED APPEAL BRIEF

Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Pursuant to 37 C.F.R. §1.192, and in response to the "Notification of Non-Compliant Appeal Brief" mailed on May 7, 2007, the Applicants hereby respectfully submit the following Amended Appeal Brief in support of their appeal. This Amended Appeal Brief replaces the original brief filed on March 19, 2007.

(1) Real Party in Interest

The real parties in interest are (a) Sony Corporation, a Japanese corporation having a primary place of business in Tokyo, Japan; and (b) Sony Electronics Inc., a U.S. corporation having a primary place of business in Park Ridge, New Jersey.

(2) Related Appeals and Interferences

No related appeals or interferences are known to the Appellant.

(3) Status of Claims

Claims 1-17, 26-30, and 32-44, which constitute the subject matter of this appeal, are pending. Claims 1-17, 26-30, and 32-44 are under final rejection. Claims 18-25, 31, and 45-56 have been previously cancelled.

(4) Status of Amendments

No amendments have been submitted subsequent to the Final Rejection in this application.

(5) Summary of Claimed Subject Matter

In the pending application, claims 1-17, 26-30, and 32-44 are pending. Claims 18-25, 31, and 45-56 have been previously cancelled. Claims 1, 6, 7, 8, 9, 10, 17, and 26 are independent claims and the remaining claims are dependent claims.

In previous systems attempts at natural language processing of human speech have been both inefficient and rigid. For example, natural language interfaces have been implemented as automated phone systems such as at airline reservation systems. Such systems prompt the user to speak within a certain context. In such systems, the received speech must be in predetermined and fixed format in order that the speech can be used by the system.

These previous approaches suffer from several disadvantages. For example, since the received speech must be in a fixed format, the use of open-ended requests, that is, requests unrestricted according to a form, format, or syntax, are unsupported in these previous systems. In fact, when an open-ended request was received, previous systems typically either ignored the request or reported an error to the user making the request.

The Applicants' invention addresses the shortcomings and limitations of previous systems. More specifically, independent claim 1 recites an interface control system for operating a plurality of devices. The system includes a 3 dimensional microphone array (e.g., arrays 108 as shown in FIG. 2 of the Application, reproduced below for the convenience of the reader) and a feature extraction module (e.g., feature extraction module 202) coupled to

the first microphone array (e.g., array 108). A speech recognition module (e.g., speech recognition module 204) is coupled to the feature extraction module (e.g., feature extraction

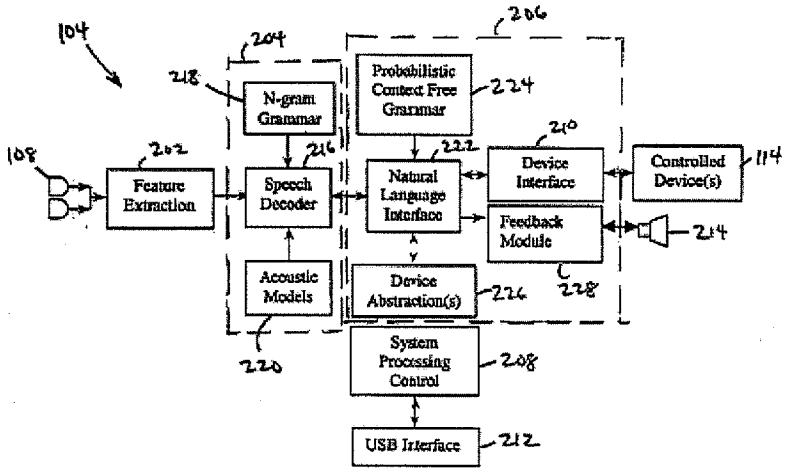

FIG. 2

module 202) and the speech recognition module utilizes hidden Markov models and can switch between different acoustic models and different grammars. Specification, page 14, lines 16- 34. In one example, these different grammars are the rules by which lexica are built with the lexica being dictionaries consisting of words and their pronunciation entries. Specification, page 19, lines 12-32. At least one of the different acoustic models and at least one of the different grammars is downloaded over a network. Specification, page 11, lines 29-34. A natural language interface module (e.g., natural language control module 206) is coupled to the speech recognition module (e.g., speech recognition module 204). A device interface (e.g., device interface 212) is coupled to the natural language interface module (e.g., natural language control module 206) and the natural language interface module operates a plurality of devices (e.g., devices 114) of one or more types that are coupled to the device interface based upon non-prompted, open-ended natural language requests from a user. Specification, page 5, line 25- page 6, line 22. The natural language interface module abstracts each of the plurality of devices into a respective one of the different grammars and a respective one of a plurality of lexica corresponding to each of the plurality of devices. Specification, page 10, line 30- page 11, line 6.

Independent claim 6 recites a natural language interface control system that operates a plurality of devices. The system includes a 3 dimensional microphone array (e.g., array 108) and a feature extraction module (e.g., feature extraction module 202) that is coupled to the first microphone array (e.g., array 108). A speech recognition module (e.g., speech recognition module 204) is coupled to the feature extraction module (e.g., feature extraction module 202) and the speech recognition module (e.g., speech recognition module 204) utilizes hidden Markov models and can switch between different acoustic models and different grammars. Specification, page 14, lines 16- 34. A natural language interface module (e.g., natural language control module 206) is coupled to the speech recognition module and a device interface (e.g., device interface 212) is coupled to the natural language interface module (e.g., natural language control module 206). The natural language interface module (e.g., natural language control module 206) operates a plurality of devices of one or more types (e.g., devices 114) that are coupled to the device interface (e.g., device interface 212) based upon non-prompted, open-ended natural language requests from a user. The natural language interface (e.g., natural language control module 206) abstracts each of the plurality of devices into a respective one of a plurality of grammars and a respective one of a plurality of lexica corresponding to each of the plurality of devices. Specification, page 10, line 30- page 11, line 6.

Independent claim 7 recites a natural language interface control system for operating a plurality of devices. The system includes a 3 dimensional microphone array (e.g., array 108) and a feature extraction module (e.g., feature extraction module 202) that is coupled to the first microphone array (e.g., array 108). A speech recognition module (e.g., speech recognition module 204) is coupled to the feature extraction module (e.g., feature extraction module 202) and the speech recognition module (e.g., speech recognition module 204) utilizes hidden Markov models and can switch between different acoustic models and different grammars. Specification, page 14, lines 16- 34. A natural language interface module (e.g., natural language control module 206) is coupled to the speech recognition module and a device interface (e.g., device interface 212) is coupled to the natural language interface module (e.g., natural language control module 206). The natural language interface module (e.g., natural language control module 206) operates a plurality of devices of one or

more types (e.g., devices 114) that are coupled to the device interface (e.g., device interface 212) based upon non-prompted, open-ended natural language requests from a user. The natural language interface module (e.g., natural language control module 206) searches for the non-prompted, open-ended user requests upon the receipt and recognition of an attention word. Specification, page 10, lines 14-29.

Independent claim 8 recites a natural language interface control system for operating a plurality of devices. The system includes a 3 dimensional microphone array (e.g., array 108) and a feature extraction module (e.g., feature extraction module 202) that is coupled to the first microphone array (e.g., array 108). A speech recognition module (e.g., speech recognition module 204) is coupled to the feature extraction module (e.g., feature extraction module 202) and the speech recognition module (e.g., speech recognition module 204) utilizes hidden Markov models and can switch between different acoustic models and different grammars. Specification, page 14, lines 16- 34. A natural language interface module (e.g., natural language control module 206) is coupled to the speech recognition module and a device interface (e.g., device interface 212) is coupled to the natural language interface module. The natural language interface module (e.g., natural language control module 206) operates a plurality of devices of one or more types (e.g., devices 114) that are coupled to the device interface (e.g., device interface 212) based upon non-prompted, open-ended natural language requests from a user. The natural language interface module (e.g., natural language control module 206) context switches grammars, acoustic models, and lexica upon receipt and recognition of an attention word. Specification, page 20, line 13- page 21, line 12.

Independent claim 9 recites a natural language interface control system for operating a plurality of devices. The system includes a 3 dimensional microphone array (e.g., array 108) and a feature extraction module (e.g., feature extraction module 202) that is coupled to the first microphone array (e.g., array 108). A speech recognition module (e.g., speech recognition module 204) is coupled to the feature extraction module (e.g., feature extraction module 202) and the speech recognition module (e.g., speech recognition module 204) utilizes hidden Markov models and can switch between different acoustic models and different grammars. Specification, page 14, lines 16- 34. A natural language interface

module (e.g., natural language control module 206) is coupled to the speech recognition module (e.g., speech recognition module 204) and a device interface (e.g., device interface 212) is coupled to the natural language interface module (e.g., natural language control module 206). The natural language interface module (e.g., natural language control module 206) operates a plurality of devices of one or more types (e.g., devices 114) that are coupled to the device interface (e.g., device interface 212) based upon non-prompted, open-ended natural language requests from a user. A grammar module (e.g., grammar module 218) stores different grammars for each of the plurality of devices.

Independent claim 10 recites a natural language interface control system for operating a plurality of devices. The system includes a 3 dimensional microphone array (e.g., array 108) and a feature extraction module (e.g., feature extraction module 202) coupled to the first microphone array (e.g., array 108). A speech recognition module (e.g., speech recognition module 204) is coupled to the feature extraction module and the speech recognition module (e.g., speech recognition module 204) utilizes hidden Markov models and can switch between different acoustic models and different grammars. Specification, page 14, lines 16- 34. A natural language interface module is coupled to the speech recognition module (e.g., speech recognition module 204) and a device interface (e.g., device interface 212) is coupled to the natural language interface module (e.g., natural language control module 206). The natural language interface module (e.g., natural language control module 206) operates a plurality of devices of one or more types (e.g., devices 114) that are coupled to the device interface (e.g., device interface 212) based upon non-prompted, open-ended natural language requests from a user. An acoustic model module (e.g., acoustic models 220) stores different acoustic models for each of the plurality of devices (e.g., devices 114).

Independent claim 17 recites searching for an attention word based on a first context including a first set of models, grammars, and lexica. Upon finding the attention word, the first context is switched to a second context to search for an open-ended user request. For example the open-ended user requests may include "I wanna watch TV", "hey, let's watch TV", "Turn on the TV", "Do you have the album 'Genesis'?" The second context includes a second set of models, grammars, and lexicons. See Specification, page 6, line 23- page 8, line 4.

Independent claim 26 recites a natural language interface control system for operating a plurality of devices. The system includes a first microphone (e.g., array 108) and a feature extraction module (e.g., feature extraction module 202) that is coupled to the first microphone (e.g., array 108). A speech recognition module (e.g., speech recognition module 204) is coupled to the feature extraction module (e.g., feature extraction module 202) and a natural language interface module (e.g., natural language control module 206) is coupled to the speech recognition module (e.g., speech recognition module 204). A device interface (e.g., device interface 212) is coupled to the natural language interface module (e.g., natural language control module 206), and the natural language interface module (e.g., natural language control module 206) operates a plurality of devices of one or more types (e.g., devices 114) that are coupled to the device interface (e.g., device interface 212) based upon non-prompted, open-ended natural language requests from a user. An external network interface is coupled to the natural language interface control system. The natural language interface (e.g., natural language control module 206) abstracts each of the plurality of devices into a respective one of a plurality of grammars and a respective one of a plurality of lexica corresponding to each of the plurality of devices. Specification, page 10, line 30- page 11, line 6.

(6) Grounds of Rejection to be Reviewed on Appeal

(A) Whether claim 17 is anticipated by U.S. Patent No. 6,584, 439 to Geilhufe?

(B) Whether claims 1-16, 26-30, and 32-44 are unpatentable under 35 U.S.C. §103 over U.S. Patent No. 6,324,512 to Junqua in view of an article by Giuliani ("Hands Free Continuous Speech Recognition in Noisy Environment Using a Four Microphone Array") and U.S. Patent No. 6,408,272 to White ("the White patent")?

(7) Argument

(A) Claim 17 is Not Anticipated by Geilhufe

As mentioned, claim 17 recites searching for an attention word based on a first context including a first set of models, grammars, and lexica. Upon finding the attention

word, the first context is switched to a second context to search for an open-ended user request. The second context includes a second set of models, grammars, and lexica.

The Examiner stated that Geilhufe teaches the processing of an open-ended request. Specifically, the Examiner stated that the phrase "Aardvark Call Mom" (used in Geilhufe) was an open-ended user request. In addition, the Examiner asserted that

> the personal name of Aardvark employs its own grammar, lexicon, and model of device names- which is inherently and undeniable a context, wherein the user must supply the word Aardvark- wherein in the context is interpreted as the application determination, secondly, the context is inherently switched to a second context (or topic) directly relating to the open ended user request, this second context employs only thereafter a second grammar, model and lexicon which it accesses after the keyword "Aardvark" is determined.

The Applicant respectfully disagrees with these statements for the reasons stated below.

As an initial matter, the Geilhufe system is not able to process open-ended requests. In fact, the requests received must be in a predetermined format or the Geilhufe system will not be able to recognize them. More specifically, Geilhufe describes a standard command syntax that is used "for ***all*** voice commands." See Geilhufe, col. 19, lines 15-37 (emphasis added). The standard syntax used in the Geilhufe system specifies that user requests must be in the form of <silence> <name> <command> <modifiers & variables>. While Geilhufe mentions two alternative command formats, only a single format is ever used. See Geilhufe, col. 19, lines 55-67 and col. 20, lines 29-36. In other words, all commands of Geilhufe must follow a fixed format and cannot deviate from the standard format, whatever that format is.

In contrast, claim 17 recites the use of open-ended commands, that is, commands that do not follow a predetermined format. For this reason, the Applicants assert that an element of claim 17 is missing from the Geilhufe reference and, consequently, claim 17 is allowable over Geilhufe.

In addition, the Geilhufe system does not switch from a first set of models, grammars, and lexica to a second set of models, grammars, and lexica upon finding an attention word as recited in claim 17. In rejecting claim 17, the Examiner analyzed the phrase "Aardvark call Mom" and asserted the phrase could be split into two portions. A first portion ("Aardvark") was deemed by the Examiner to be an attention word and a second portion ("call mom") was

deemed to be an open-ended command. Furthermore, the Examiner stated that each of the two portions "inherently" employed a separate grammar, models, and lexica. The Applicants disagree with these statements for the reasons stated below.

Specifically, the Examiner's assertions are contradicted by the express teachings of Geilhufe. For instance, according to Geilhufe, ***an entire phrase*** is analyzed according to a ***single*** syntax. See Geilhufe, col. 19, lines 15-17. For instance, all received speech commands are analyzed according to a single syntax (e.g., <silence> <name> <command> <modifiers & variables>). *Id.* There is simply no teaching or suggestion in Geilhufe that any received speech phrase is split into separate portions and each of the portions is analyzed according to a different grammar, model, or lexicon[1].

Even if the Geilhufe system were to split a speech phrase into multiple portions and each portion were separately analyzed, there is no "inherent" reason why each portion would need to be analyzed according to a separate grammar, lexicon, or model. To take one example, since a lexicon is typically a dictionary of words and their pronunciation entries, a single lexicon could be used, especially when the devices were similar in type. And, in another example, it is possible that a single lexicon is used when it was necessary to conserve memory space. In other words, the Applicants assert the Geilhufe fails to teach that a phrase is split into separate portions and, even if it did, Geihufe fails to teach that these separate portions would be required to be analyzed according to different grammars, lexica, or models.

Since Geilhufe fails to teach or suggest switching between first and second grammars, models, and lexica, as is recited in claim 17, the Applicants assert that claim 17 is not anticipated by Geilhufe.

(B) Claims 1-16, 26-30, and 32-44 are not unpatentable over Junqua in view of Giuliani and White

As mentioned, claim 1 recites an interface control system for operating a plurality of devices. The system includes a 3 dimensional microphone array and a feature extraction

[1] There would also be no advantage or motivation to modify the Geilhufe system in order to use different grammars, lexica, and models in order to analyze different portions of the same phrase. To the contrary, such an approach would greatly complicate the processing that occurs within the Geilhufe system.

module coupled to the first microphone array. A speech recognition module is coupled to the feature extraction module and the speech recognition module utilizes hidden Markov models and can switch between different acoustic models and different grammars. At least one of the different acoustic models and at least one of the different grammars is downloaded over a network. A natural language interface module is coupled to the speech recognition module. A device interface is coupled to the natural language interface module and the natural language interface module operates a plurality of devices of one or more types that are coupled to the device interface based upon non-prompted, open-ended natural language requests from a user. The natural language interface module abstracts each of the plurality of devices into a respective one of the different grammars and a respective one of a plurality of lexica corresponding to each of the plurality of devices.

The Office Action stated that Junqua teaches all of the elements of the claim "but lacks explicitly wherein the natural language interface module abstracts each of the plurality of devices into a respective one of the different grammars and a respective one of a plurality of lexica corresponding to each of the plurality of devices." However, the Office Action stated that

> Geilhufe teaches an interface that abstracts...each of the plurality of devices (C.17.lines 6-10, C.19.lines 33-37, C.18.lines 1-4-wherein each device has "abstracted", core commands, and commands specific to a given application... it would have been obvious to modify Junqua's natural language parser and unified access controller with Geilhufe's device specific grammar and lexicon (vocabulary /specific list of commands). The motivation for doing so would have been to each device respond to specific commands appropriately (C.18.lines 1-4, 47-57- wherein "Aardvark call mom" results in calling mom from a desktop phone, by a command definition of a call as a specific command to a phone device, and, not, for example, a transcription of "Aardvark call mom" into a document.

The Applicants respectfully disagree with these assertions for the reasons stated below.

More specifically, Geilhufe does not teach abstracting each of the device types into different grammars and lexica as is recited in claim 1. In fact, Geilhufe teaches the use of a single syntax as has been described above.

Additionally, Geilhufe does not teach or suggest the use "device specific" grammars as asserted by the Examiner. The single syntax is not related or associated with a particular type of device. Instead, the single syntax of commands in the Geilhufe system must work with ***all*** device types.

In addition, as mentioned above, Geilhufe is not able to process open-ended user requests as is recited in claim 1. As mentioned previously, the commands in Geilhufe must be in a fixed format.

Consequently, the Applicants assert that claim 1 is allowable over the proposed combination since Geilhufe fails to teach or suggest the above-mentioned claim elements.

Furthermore, even assuming these elements were somehow taught by Geilhufe, there is no suggestion or motivation to modify Junqua to have a natural language interface module that abstracts each of the plurality of devices into a respective one of the different grammars and a respective one of a plurality of lexica corresponding to each of the plurality of devices. There must be a motivation to make the proposed modification either in the references themselves or apparent to one skilled in the art. See MPEP § 2143.01.

More specifically, Junqua teaches the use of a tuner 40 and a recorder 44 that are activated by a controller module 30. See FIG. 1 of Junqua, reproduced below for the convenience of the reader.

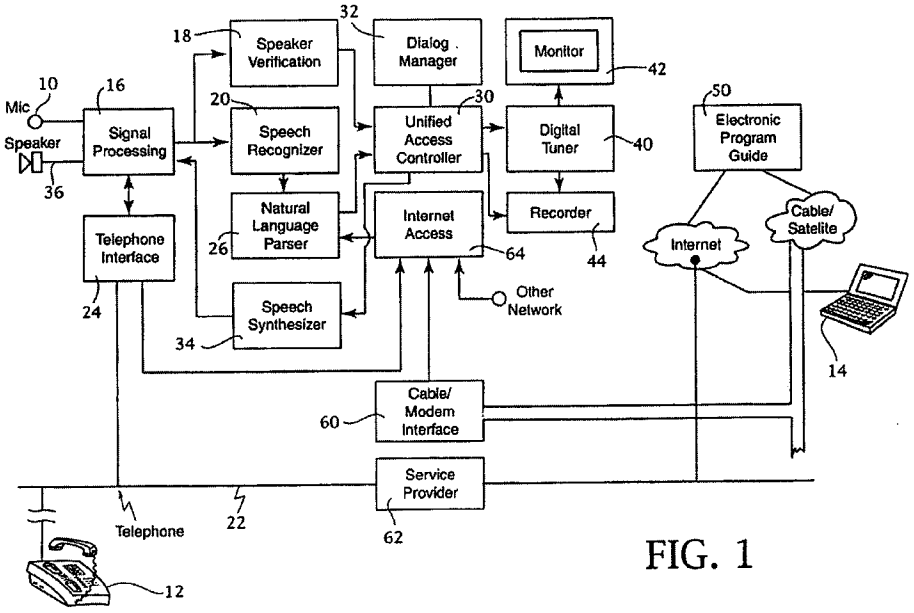

FIG. 1

The user's spoken instructions are converted into text by speech recognizer 20 and the output of the speech recognizer 20 is supplied to the natural language parser 26. Junqua, col. 2, lines 52-61. The natural language parser 26 then parses the text. *Id.* The output of the parser 26 is sent to the controller module 30, which sends electrical signals to activate the tuner 40 and the recorder 44. In other words, the architecture of the Junqua system mandates that the controller module 30 receives information in a standard text format.

To modify the parser module 26 to be able to abstract each of the plurality of devices into a different grammar and lexicon that correspond to the plurality of devices of one or more types is simply not suggested in the reference. In fact, the devices controlled by the control module 30 in the Junqua system are all of the same or similar type (i.e., relating to video or video control devices). As can be seen in FIG. 1 of Junqua, different types of devices such as the telephone 12 and computer 14 are not controlled by the controller module 30. Consequently, to achieve uniformity and design efficiency, one skilled in the art would

be motivated to use the ***same*** grammar and lexicon for each device rather than a different grammar and a different lexicon for the same type of devices. For instance, since the spoken words used to control each of the devices 40 and 44 are likely to be similar (both relating to video technology), the same lexicon would likely be used.

It is not proper for the Examiner to scour the prior art to find the elements of the Applicants' claims and then use the Applicants' own teachings to combine these elements as claimed. See MPEP § 2143.01. Consequently, the Applicants assert that the proposed modification is non-obvious and the Applicants assert that claim 1 is allowable for this additional reason.

Claims 6, 7, 8, 9, 10, and 26 are independent claims that recite the use of open-ended requests and the switching between different grammars, lexicon, and models. Consequently, the Applicants assert that claims 6, 7, 8, 9, 10, and 26 are allowable for the same reasons as described above with respect to claim 1.

Claims 2-5, 11-15, 27-30, and 32-44 ultimately depend upon claims 1, 6, 7, 8, 9, 10, and 26, which have been shown to be allowable above, and therefore, these claims are also allowable. In addition, they introduce additional content that, particularly when considered in context with the claim from which they depend, introduce additional incremental patentable subject matter. Accordingly, the Applicants reserve the right to present further arguments in the future with regard to this dependent claim if independent claims 1, 6, 7, 8, 9, 10, and 26 are found to be unpatentable.

(8) Claims Appendix

Claim 1 (Previously presented): A natural language interface control system for operating a plurality of devices comprising:

a 3 dimensional microphone array;

a feature extraction module coupled to the first microphone array;

a speech recognition module coupled to the feature extraction module, wherein the speech recognition module utilizes hidden Markov models and can switch between different acoustic models and different grammars, wherein at least one of the different acoustic models and at least one of the different grammars is downloaded over a network;

a natural language interface module coupled to the speech recognition module; and

a device interface coupled to the natural language interface module, wherein the natural language interface module is for operating a plurality of devices of one or more types that are coupled to the device interface based upon non-prompted, open-ended natural language requests from a user;

wherein the natural language interface module abstracts each of the plurality of devices into a respective one of the different grammars and a respective one of a plurality of lexica corresponding to each of the plurality of devices.

Claim 2 (Original): The system of Claim 1 further comprising the plurality of devices coupled to the natural language interface module.

Claim 3 (Original): The system of Claim 1 wherein the speech recognition module utilizes an N gram grammar.

Claim 4 (Original): The system of Claim 1 wherein the natural language interface module utilizes a probabilistic context free grammar.

Claim 5 (Previously presented): The system of Claim 1 wherein the microphone array comprises said 3 dimensional microphone array further comprising a planar microphone array and at least one linear microphone array located in a different plane in space.

Claim 6 (Previously presented): A natural language interface control system for operating a plurality of devices comprising:

a 3 dimensional microphone array;

a feature extraction module coupled to the first microphone array;

a speech recognition module coupled to the feature extraction module, wherein the speech recognition module utilizes hidden Markov models and can switch between different acoustic models and different grammars;

a natural language interface module coupled to the speech recognition module; and

a device interface coupled to the natural language interface module, wherein the natural language interface module is for operating a plurality of devices of one or more types that are coupled to the device interface based upon non-prompted, open-ended natural language requests from a user;

wherein the natural language interface abstracts each of the plurality of devices into a respective one of a plurality of grammars and a respective one of a plurality of lexica corresponding to each of the plurality of devices.

Claim 7 (Previously presented): A natural language interface control system for operating a plurality of devices comprising:

a 3 dimensional microphone array;

a feature extraction module coupled to the first microphone array;

a speech recognition module coupled to the feature extraction module, wherein the speech recognition module utilizes hidden Markov models and can switch between different acoustic models and different grammars;

a natural language interface module coupled to the speech recognition module; and

a device interface coupled to the natural language interface module, wherein the natural language interface module is for operating a plurality of devices of one or more types that are coupled to the device interface based upon non-prompted, open-ended natural language requests from a user;

wherein the natural language interface module searches for the non-prompted, open-ended user requests upon the receipt and recognition of an attention word.

Claim 8 (Previously presented): A natural language interface control system for operating a plurality of devices comprising:

a 3 dimensional microphone array;

a feature extraction module coupled to the first microphone array;

a speech recognition module coupled to the feature extraction module, wherein the speech recognition module utilizes hidden Markov models and can switch between different acoustic models and different grammars;

a natural language interface module coupled to the speech recognition module; and

a device interface coupled to the natural language interface module, wherein the natural language interface module is for operating a plurality of devices of one or more types that are coupled to the device interface based upon non-prompted, open-ended natural language requests from a user;

wherein the natural language interface module context switches grammars, acoustic models, and lexica upon receipt and recognition of an attention word.

Claim 9 (Previously presented): A natural language interface control system for operating a plurality of devices comprising:

a 3 dimensional microphone array;

a feature extraction module coupled to the first microphone array;

a speech recognition module coupled to the feature extraction module, wherein the speech recognition module utilizes hidden Markov models and can switch between different acoustic models and different grammars;

a natural language interface module coupled to the speech recognition module;

a device interface coupled to the natural language interface module, wherein the natural language interface module is for operating a plurality of devices of one or more types that are coupled to the device interface based upon non-prompted, open-ended natural language requests from a user; and

a grammar module for storing different grammars for each of the plurality of devices.

Claim 10 (Previously presented): A natural language interface control system for operating a plurality of devices comprising:

a 3 dimensional microphone array;

a feature extraction module coupled to the first microphone array;

a speech recognition module coupled to the feature extraction module, wherein the speech recognition module utilizes hidden Markov models and can switch between different acoustic models and different grammars;

a natural language interface module coupled to the speech recognition module;

a device interface coupled to the natural language interface module, wherein the natural language interface module is for operating a plurality of devices of one or more types that are coupled to the device interface based upon non-prompted, open-ended natural language requests from a user; and

an acoustic model module for storing different acoustic models for each of the plurality of devices.

Claim 11 (Original): The system of Claim 1 wherein the device interface comprises a wireless device interface.

Claim 12 (Original): The system of Claim 1 further comprising an external network interface coupled to the natural language interface control system.

Claim 13 (Original): The system of Claim 1 further comprising a remote unit containing a first microphone array, the feature extraction module, the speech recognition module, and the natural language interface module, wherein said 3 dimensional microphone array includes the first microphone array.

Claim 14 (Original): The system of Claim 13 further comprising a base unit coupled to the remote unit.

Claim 15 (Previously presented): The system of Claim 14 wherein the base unit includes a second microphone array, wherein said 3 dimensional microphone array includes the second microphone array.

Claim 16 (Previously presented): The system of Claim 15 wherein the first microphone array and the second microphone array implement said 3 dimensional microphone array.

Claim 17 (Previously presented): A method of speech recognition comprising:

searching for an attention word based on a first context including a first set of models, grammars, and lexica; and

switching, upon finding the attention word, to a second context to search for an open-ended user request, wherein the second context includes a second set of models, grammars, and lexicons.

Claims 18-25 (Canceled)

Claim 26 (Previously presented): A natural language interface control system for operating a plurality of devices comprising:

a first microphone;

a feature extraction module coupled to the first microphone;

a speech recognition module coupled to the feature extraction module;

a natural language interface module coupled to the speech recognition module;

a device interface coupled to the natural language interface module, wherein the natural language interface module is for operating a plurality of devices of one or more types that are coupled to the device interface based upon non-prompted, open-ended natural language requests from a user; and

an external network interface coupled to the natural language interface control system;

wherein the natural language interface abstracts each of the plurality of devices into a respective one of a plurality of grammars and a respective one of a plurality of lexica corresponding to each of the plurality of devices.

Claim 27 (Previously presented): The system of Claim 26 further comprising the plurality of devices coupled to the natural language interface module.

Claim 28 (Previously presented): The system of Claim 26 wherein the speech recognition module utilizes an N gram grammar.

Claim 29 (Previously presented): The system of Claim 26 wherein the natural language interface module utilizes a probabilistic context free grammar.

Claim 30 (Previously presented): The system of Claim 26 wherein the microphone array comprises a 3 dimensional microphone array comprising a planar microphone array and at least one linear microphone array located in a different plane in space.

Claim 31 (Canceled)

Claim 32 (Previously presented): The system of Claim 26 wherein the natural language interface module searches for the non-prompted, open-ended user requests upon the receipt and recognition of an attention word.

Claim 33 (Previously presented): The system of Claim 26 wherein the natural language interface module context switches grammars, acoustic models, and lexica upon receipt and recognition of an attention word.

Claim 34 (Previously presented): The system of Claim 26 further comprising a grammar module for storing different grammars for each of the plurality of devices.

Claim 35 (Previously presented): The system of Claim 26 further comprising an acoustic model module for storing different acoustic models for each of the plurality of devices.

Claim 36 (Previously presented): The system of Claim 26 wherein the device interface comprises a wireless device interface.

Claim 37 (Previously presented): The system of Claim 26 further comprising a remote unit containing the first microphone array, the feature extraction module, the speech recognition module, and the natural language interface module.

Claim 38 (Previously presented): The system of Claim 37 further comprising a base unit coupled to the remote unit.

Claim 39 (Previously presented): The system of Claim 38 wherein the base unit includes a second microphone array.

Claim 40 (Previously presented): The system of Claim 39 wherein the first microphone comprises a first microphone array, and said first microphone array and the second microphone array implement a 3 dimensional microphone array.

Claim 41 (Previously presented): The system of Claim 26 further comprising a central database coupled to said external network interface, said central database including at least one of grammars; speech models; device abstractions; programming information; and lexica.

Claim 42 (Previously presented): The system of Claim 41 wherein said central database is coupled to said external network interface through an external network.

Claim 43 (Previously presented): The system of Claim 42 further comprising:
a remote server coupled to said external network and to
said central database.

Claim 44 (Previously presented): The system of Claim 42 further comprising:
another natural language interface control system; and
another external network interface coupled to said other natural language interface control system, and to said external network.

Claims 45-56 (Canceled)

(9) Evidence Appendix

Not Applicable.

(10) Related Proceedings Appendix

Not applicable.

In view of the foregoing, it is submitted that the application is in condition for allowance which is respectfully requested. The Commissioner is hereby authorized to charge any additional fees which may be required to Deposit Account No. 06-1135.

Respectfully submitted,

FITCH, EVEN, TABIN & FLANNERY

By: Timothy R. Baumann
Timothy R. Baumann
Registration No. 40,502

Date: 6/7/2007

Suite 1600
120 South LaSalle
Chicago, Illinois 60603-3406
Telephone: (312) 577-7000
Facsimile: (312) 577-7007